ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

BENCHMARKING OF A REGULATORY URBAN AIR QUALITY MODELLING SYSTEM UTILIZING PROBENCH

..first impressions by

Ari Karppinen ***(AQ dispersion modeling group)***

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Contents

- **Introduction**

- **Short history of evaluation exercises @ FMI**
- **Complete harmonization of evaluation procedures - is it ever possible?**
- **Examples & first impressions with Delta-tool/PROBENCH**

- **Conclusions**

Introduction *(= my personal guidelines from the early 90's)*

Components of model evaluation

(Steven Hanna, 1991)

1. Scientific review
2. Code fidelity to equations
3. Statistical evaluation with experimental data
4. Model to model comparison
5. Sensitivity study
6. Ease of use

Introduction

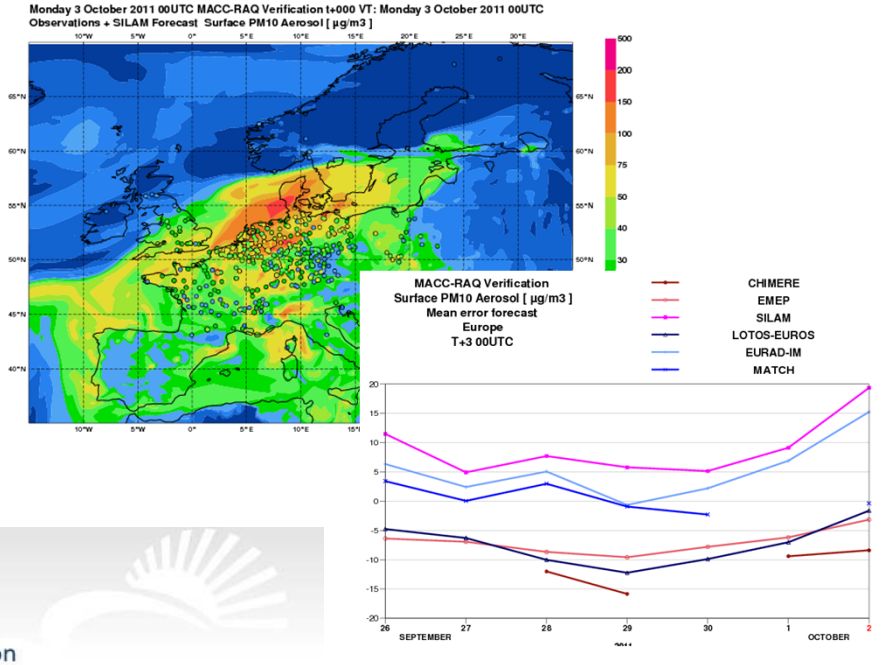
- "Models are validated with measured values, not actual values!"
- "Specific events are not predictable, only the set of possible events can be described."
- "All models are wrong, but some of them are useful."
- "Models can not be validated, they only can be invalidated."

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Evaluation/verification still key issue for the work - some ongoing projects :

Monitoring atmospheric composition & climate

http://macc-raq.gmes-atmosphere.eu/



PESCaDO
Personalized Environmental Service Configuration and Delivery Orchestration

FMI-role: develop/implement methods for dynamical ensembling of environmental information (web-services)

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

”Long” history of evaluation studies (~20 yrs)

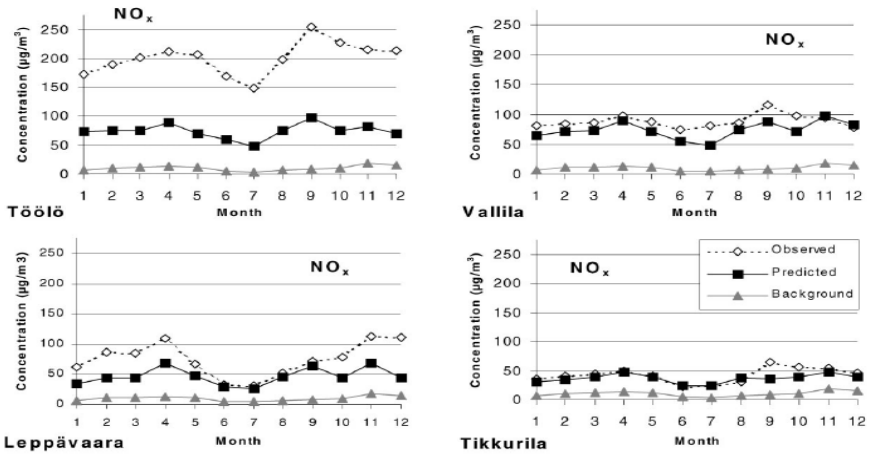


Fig. 2. a–d The predicted and measured monthly averages of the NO_x concentration (µg m^{-3}) at the four monitoring stations considered, together with the regional background concentrations in 1993.

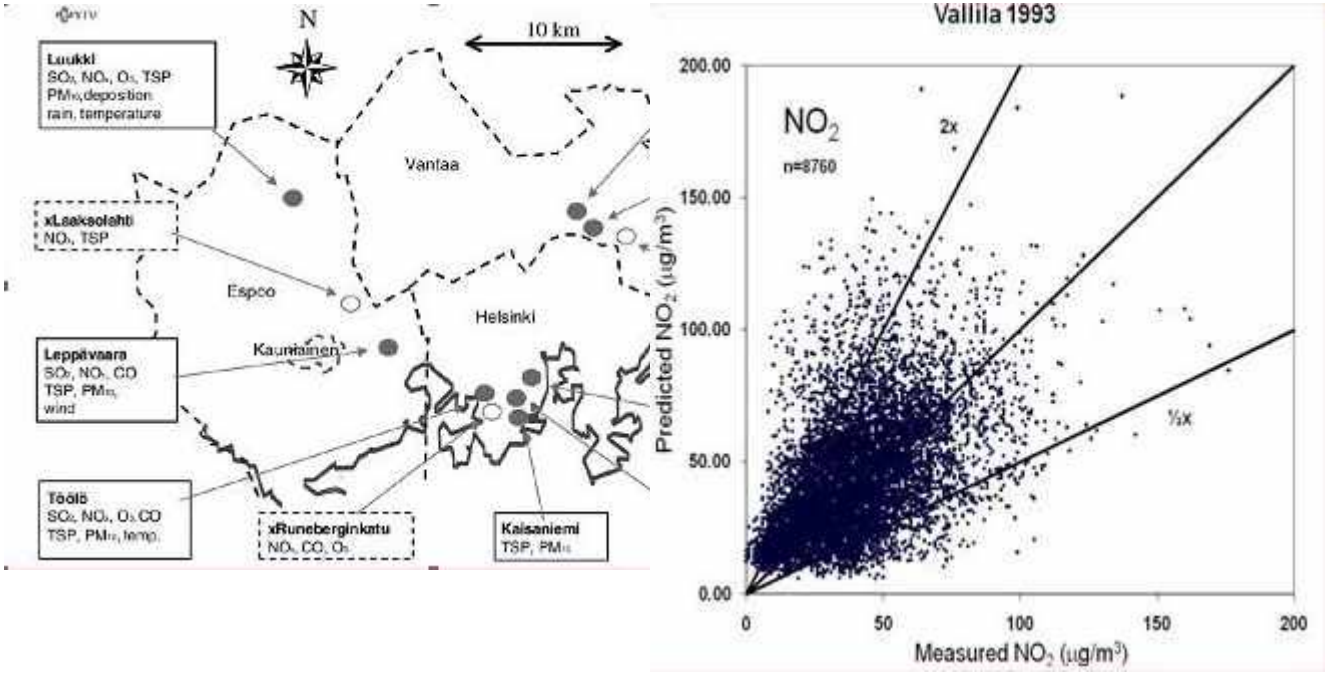
Perusaineisto												
Tammikuu 1993												
Tuntiarvot				Töölö SO2	Töölö NO	TöölöNO2	Töölö CO	Töölö O3	ValliSO2	VallilNO	ValliNO2	VallilCO
Vv	Kk	Pv	HH	µg/m3	µg/m3	µg/m3	mg/m3	µg/m3	µg/m3	µg/m3	µg/m3	mg/m3
93	1	1	0	2	12	11	0,5	52	0	2	21	0,2
93	1	1	1	2	16	16	0,8	50	1	4	26	0,2
93	1	1	2	2	12	10	0,6	56	2	1	25	0,2
93	1	1	3	2	7	9	0,5	61	2	0	17	0,2
93	1	1	4	2	6	8	0,4	63	3	0	17	0,2
93	1	1	5	2	1	6	0,3	66	2	0	12	0,2
93	1	1	6	2	2	4	0,3	65	1	0	8	0,3
93	1	1	7	1	4	6	0,4	63		0	8	0,3
93	1	1	8	1	6	7	0,4	63	1	0	11	0,3
93	1	1	9	1	8	9	0,3	62	1	0	14	0,3
93	1	1	10	1	10	7	0,3	62	0	0	15	0,3
93	1	1	11	1	8	8	0,2	60	0	0	15	0,2
93	1	1	12	2	13	12	0,2	55	3	5	27	0,1

~1993->

The statistical analysis of the predicted and measured hourly time series of NO_2 concentrations in 1993 (the total number of data $N = 8760$). The statistical parameters have been defined in the text

NO_2	Töölö		Vallila		Leppävaara		Tikkurila	
Statistical parameter	Predict.	Measured	Predict.	Measured	Predict.	Measured	Predict.	Measured
Mean (µg m^{-3})	43	45	44	40	31	34	27	24
Maximum (µg m^{-3})	191	167	211	176	179	129	144	171
Stand. dev. (µg m^{-3})	22	26	23	21	21	21	18	18
Index of agreement	− 0.75		0.69		0.73		0.79	
Correlation coeff.	0.57		0.50		0.54		0.65	
Normalised MSE	0.28		0.26		0.45		0.33	
Fractional bias	− 0.045		0.095		− 0.092		0.118	

”Long” history of evaluation studies (~20 yrs)

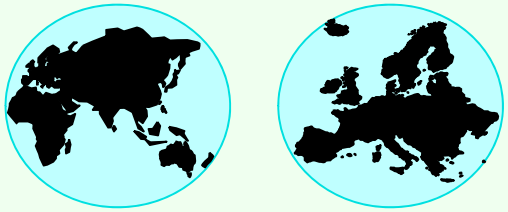


ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Weather prediction models

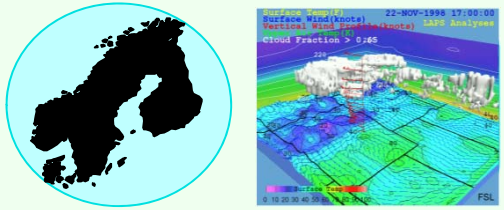
ECMWF

HIRLAM RCR



HIRLAM MBE

AROME, LAPS



Dispersion models - long-range, regional

SILAM LRT, meso-scale, radioactivity

HILATAR LRT, meso-scale

Aerosol process models:
UHMA (U Helsinki, FMI)
MONO32 (U Helsinki, Stadia)
SALSA(UH,UKU,FMI)->
SILAM-APM (FMI)

Dispersion and effects models – urban, local

CAR-FMI, roadside

UDM-FMI, urban

OSPM (NERI), street canyon

BUOYANT, fires

ESCAPE, chemical accidents

FINFLO,FLUENT(CFD) ,PALM, DNS-code development

EXPAND (FMI, YTV) population exposure

MPP-FMI, Meteorological pre-processing model

FMI-modeling system

Some application areas

- **Local/urban scale (UDM-FMI/80's)**
- **Traffic (CAR-FMI/90's)**
- **Hazardous substances (ESCAPE/90´s)**
- **Reg.scale : N/S deposition (HILATAR/90's)**
- **Regional scale (SILAM/late 90's->20**)**
  - Nuclear accidents
  - Pollen
  - Forest fires
  - Particulate matter
  - General regional scale AQ

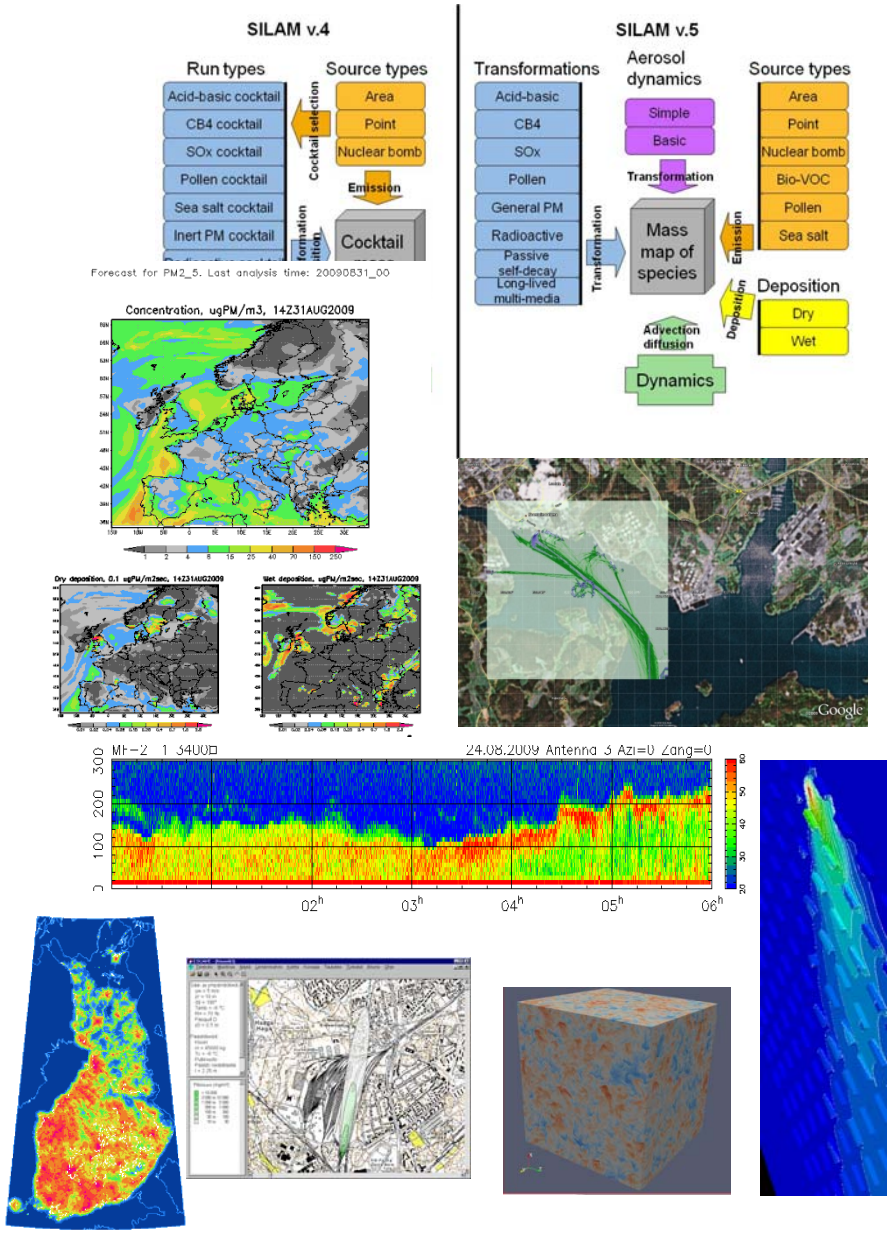
- **Microscale/CFD/LES modeling (DNS,PALM..)(10's)**

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Multitude of different types of models fit for specific purposes– can we really harmonize the evaluation procedure for all of them

–

even just inside FMI ?



(not so definite) Answer:

- **At the moment we are very far from it !**
- **One of the most critical questions : the availability of data suitable for evaluation !**
- **Standard AQ-monitoring data (<10 points for whole Helsinki area) can be used for ”simple” regulatory model evaluation – but not necessary useful for any other purpose =>**
- **New , <u>better tools</u> will help in harmonizing the evaluation procedure - but still a lot of (continuous) work has to be done to provide good quality evaluation data : strong co-operation (European, world wide) carefully planned campaigns, wind-tunnel studies, etc..)**

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

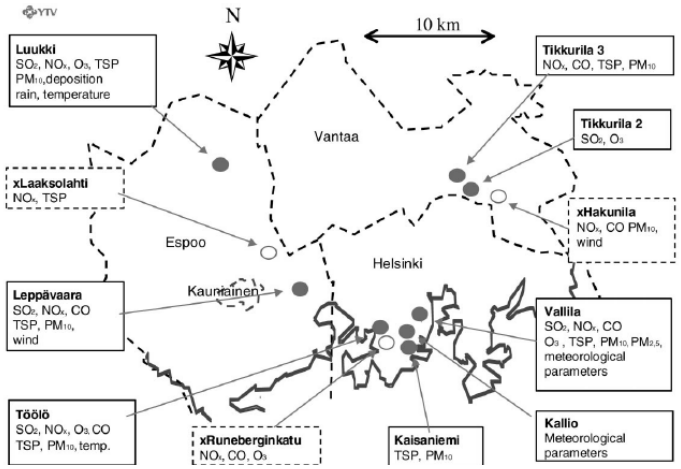
Materials for the DELTA-tool experiments

Available regulatory modelling data (Helsinki area):

"database" of predicted and measured hourly air quality data during 1993-2010

- **air quality model results for 10 full years during the period *(modelled components vary)***
- **measurement data for 7 monitoring stations in Helsinki metropolitan Area for ~ 18 years**

(locations of stations vary)



The ONLY EQ-slide……

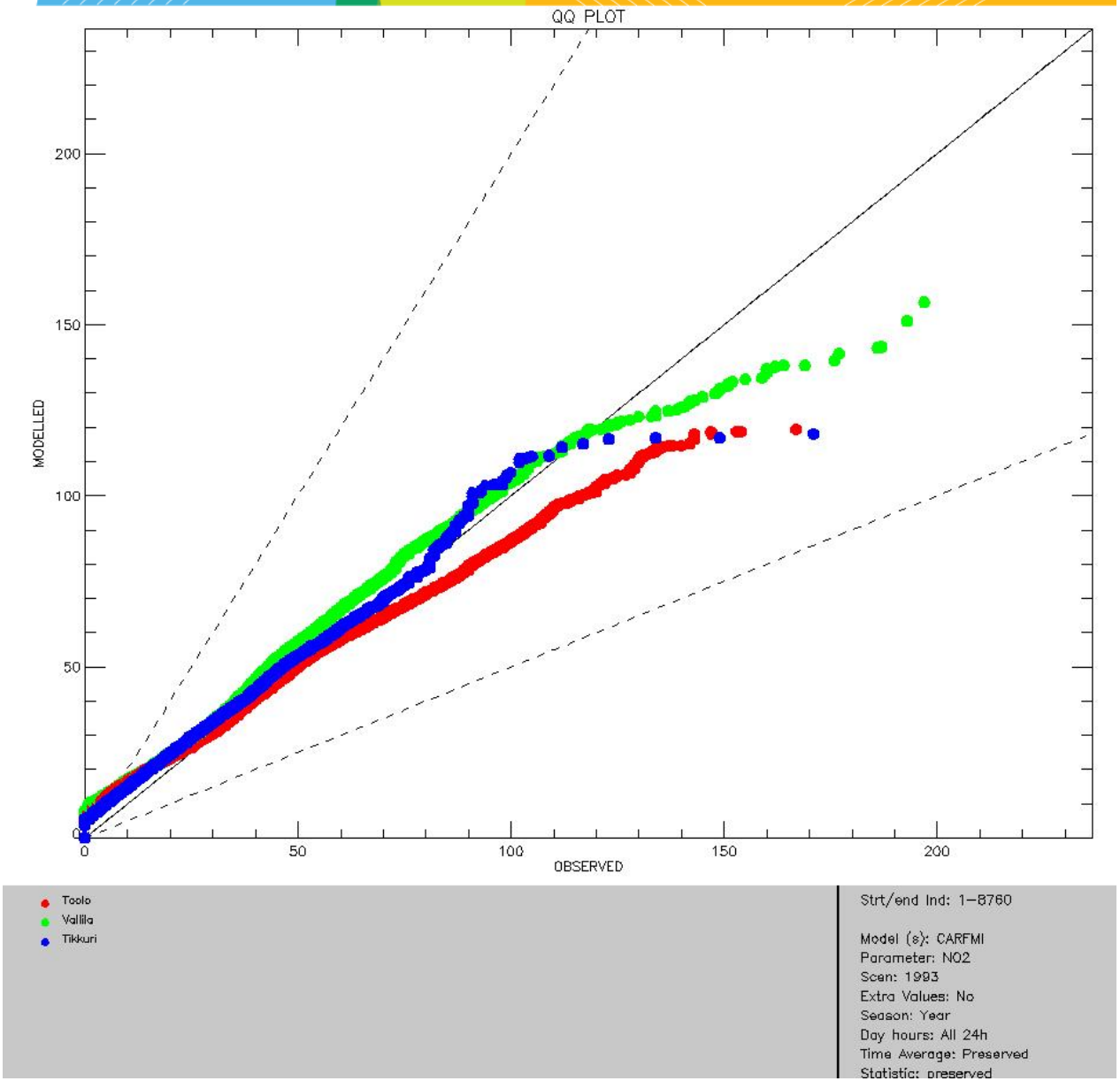
RootMeanSquare Error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2}$
Target	$RMSE / \sigma_O = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2} \Big/ \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - \overline{O})^2}$
Centred Root Mean Square error	$CRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[(M_i - \overline{M}) - (O_i - \overline{O})\right]^2}$
Model Efficiency Score	$MEF = 1 - RMSE^2$
Mean Bias	$MBias = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

DELTA: 1

Basic QQ-plots for model evaluation

Note! *Included in general model evaluation guidelines for decades*



ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

DELTA 2:

Detailed (visual) view on the performance of models,

Shows clearly which statistical indicators show acceptable model skill

Points clearly out if the model skill is consistent at all different monitoring locations

SUMMARY STATISTICS

Nb of stations: 3 valid / 3 selected

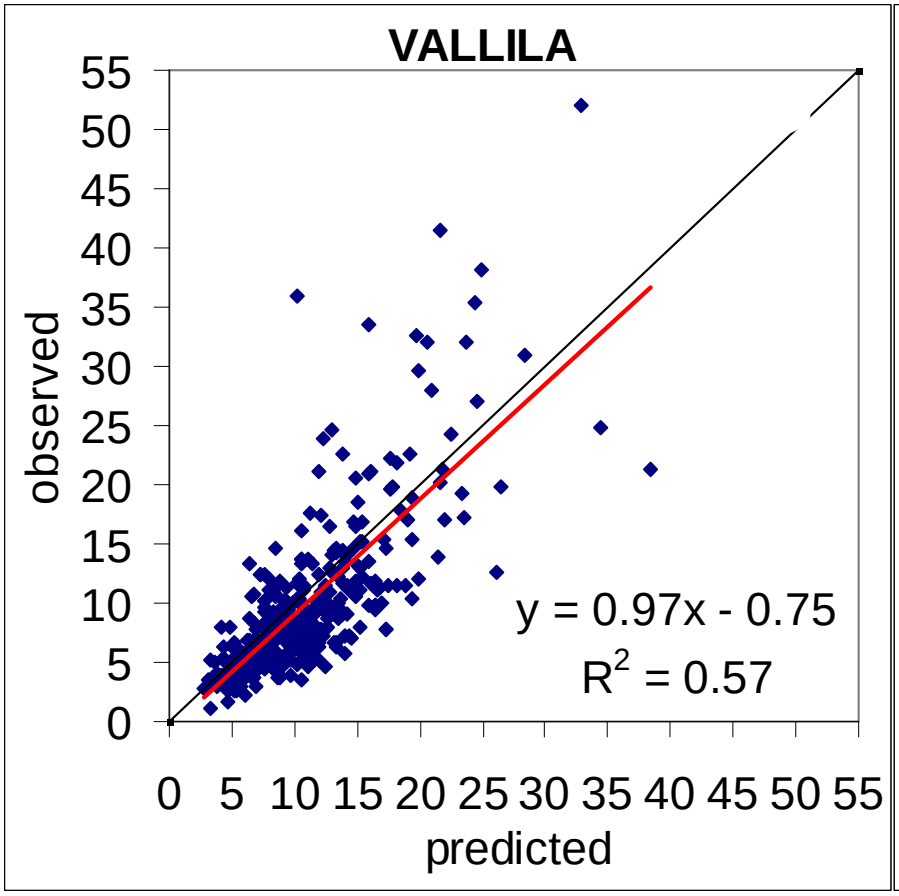
INDIC	(Crit - Goal)	90% percentile	Min	Mean	Max
TARGET	(1.00–0.80)	P 1.12 G	0.87	0.97	1.12
MFBI	(0.30–0.15)	P 0.26 G	0.06	0.16	0.26
R	(0.55–0.65)	P 0.38 G	0.38	0.50	0.57
FAC2	(0.50–0.80)	P 0.67 G	0.67	0.72	0.75
SFBI	(0.50–0.40)	P 0.21 G	0.01	0.09	0.21
RDE	(0.50–0.42)	P 0.26	0.20	0.23	0.26
RPE	(0.50–0.42)	P 0.15	0.05	0.10	0.15

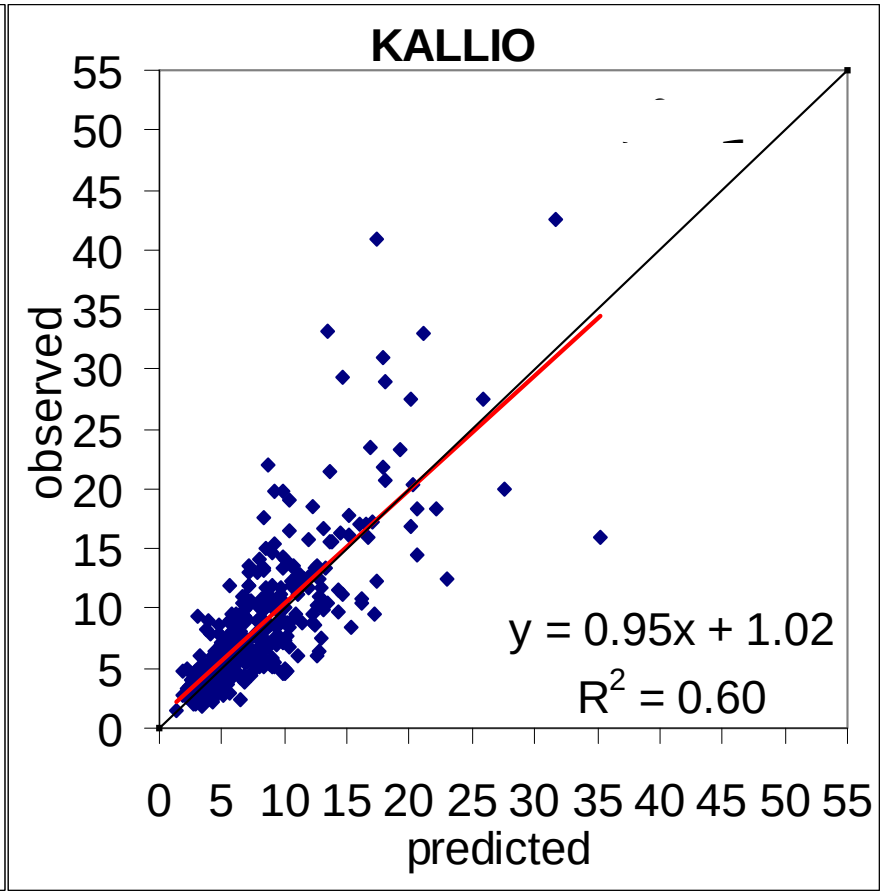
Strt/end Ind: 1–8760

Model (s): CARFMI
Parameter: NO2
Scen: 1993
Extra Values: No
Season: Year
Day hours: All 24h
Time Average: Preserved

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Predicted vs. observed daily mean
$PM_{2.5}$ concentrations at two stations – scatter plot, Correlation Coefficient squared (R^2) and Index of Agreement (IA)





VALLILA: R^2 = 0.57, IA = 0.84

KALLIO: R^2 = 0.60, IA = 0.86

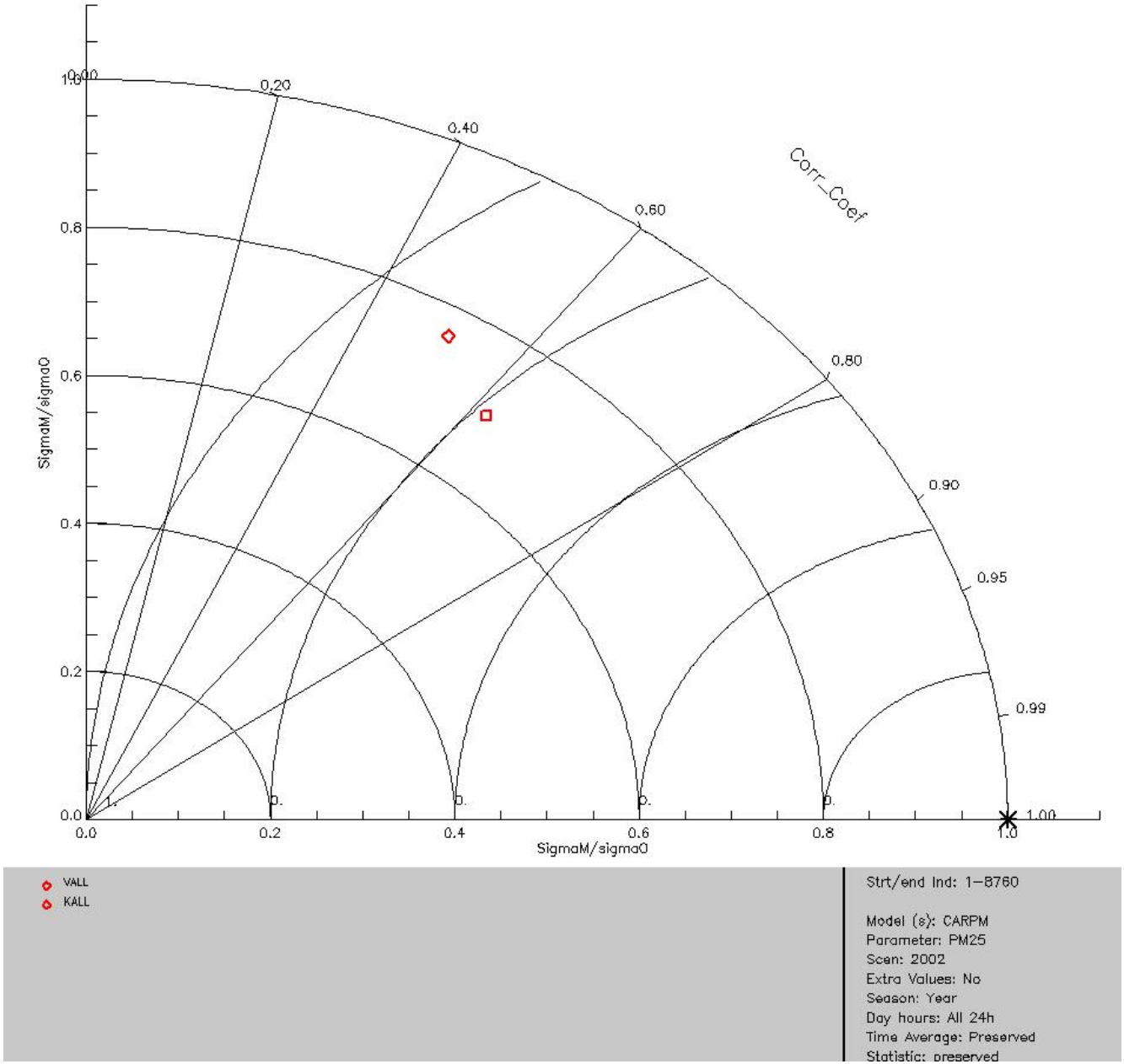
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

DELTA3:

Taylor diagram

correlation coefficient,
standard deviation
centered RMSE

Good addition to
visual
model skill
assessment



PROBENCH: Important suggestion:

MAIN statistical indicator:

distance from origin on the Target diagram
= normalized RMSE

=> ”better than average” model performance
+ list of other ”good” indicators for model skill
guaranteed if the criteria is passed

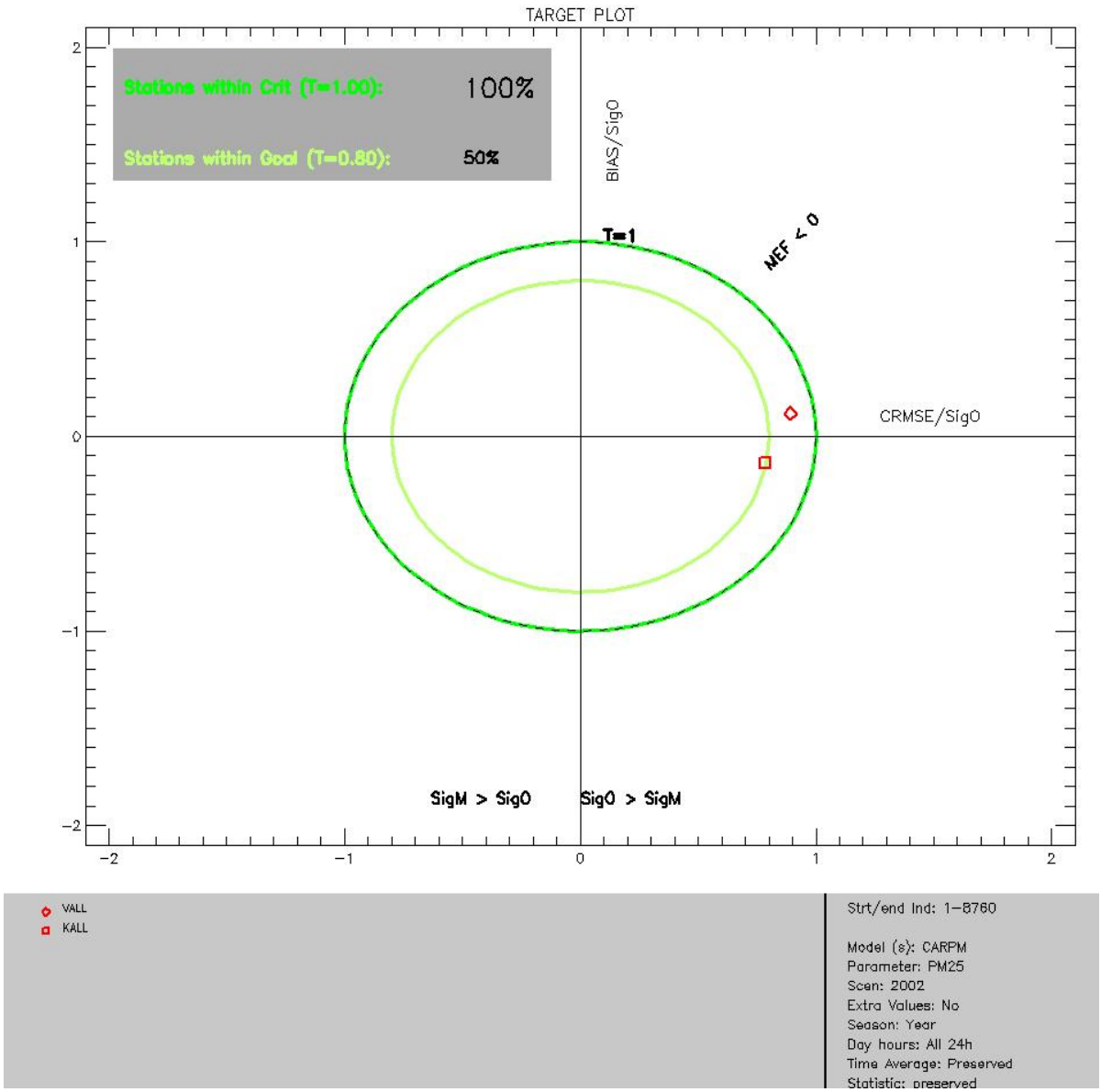
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

DELTA 4:

TARGET diagram

R, MBias,
RMSE/sigma(obs)
MEF
sign(RMSEs/RMSEo)

Something completly new (for us)…



ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

CONCLUSIONS/general

The basic statistical parameter set utilized in the FMI AQM-evaluations has been quite consistent during the last 15 years - obviously the statistical parameters used are only a subset of the parameters provided by DELTA tool.

Earlier evaluations did not clearly specify any clear acceptance criteria for the model behaviour, while PROBENCH /DELTA has added a **multitude of statistical indicators + suggestion for the main statistical indicator with clear acceptance levels** to help to identify unacceptable model performance.

Conclusions/general

For each statistical indicator, two quality bounds are proposed:

- a performance criterion which states whether sufficient quality for policy application is reached
- a performance goal which points to the optimum quality level that a model is expected to reach.

These two quality bounds **will greatly assist the user** in assessing the quality of the model performances for a given *regulatory* AQM application.

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Conclusions/re-evaluation

DELTA tool has been utilized to re-assess model calculations for Helsinki Area based on archived model and measurement data for the area (work ongoing..only part of the complete “database” has been processed)

New evaluation did not (yet) reveal any major discrepancies

..but at least the new concept and tool already proved to be applicable tool for processing data created long time before any format or content requirements for DELTA tool were created.

Some practical “issues” experienced but no major problems observed

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Conclusions/practical issues

some investments (=time and resources) are always needed to learn to use a new tool effectively – DELTA is no exception on this rule

First experiences with the tool are promising , and *seem to* prove, that the resources invested are going to be returned in the future

CONCLUSIONS

Assuming that a continuous support and resources for further development for the tool can be guaranteed, it will be an important step towards harmonizing the European model evaluation practices in the future
(=just my humble personal opinion)

neither DELTA nor any other tool can completely cover and automate the complicated process of model evaluation => the skills & experience of the evaluator AND availability of applicable evaluation data (extensive measurement campaigns/wind tunnel studies ets..) will even in the future be the decisive factor

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Thank you for your attention !

http://fairmode.ew.eea.europa.eu/models-benchmarking-sg4/delta_tool_concepts_userguide.pdf

FAIRMODE
Forum for air quality modelling in Europe

More info on the FMI evaluation ”history”

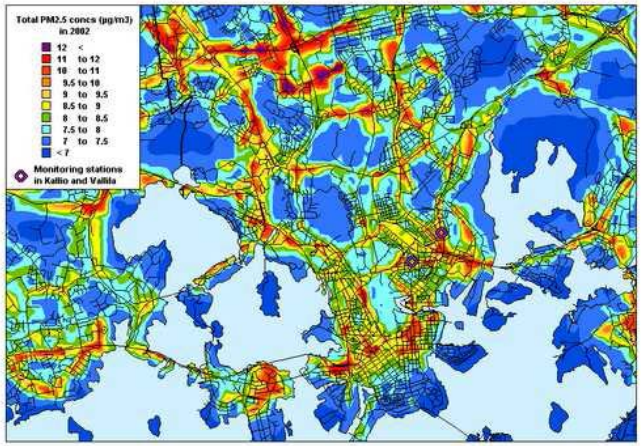
GOOGLE ! (scopus/..):

’karppinen’ ’fmi’ ’evaluation’

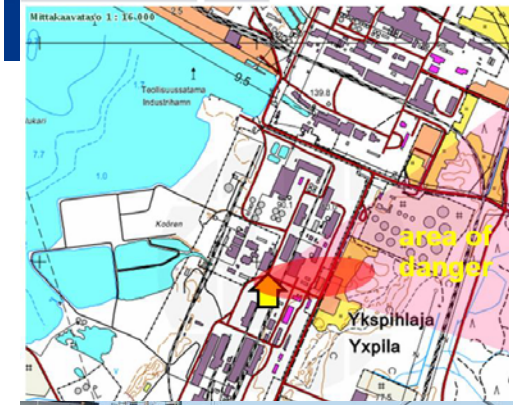
Mean	$\overline{M} = \frac{1}{N}\sum_{i=1}^{N} M_i \ , \ \overline{O} = \frac{1}{N}\sum_{i=1}^{N} O_i$
Standard Deviation	$\sigma_M = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - \overline{M})^2} \ , \ \sigma_O = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - \overline{O})^2}$
Mean Bias	$MBias = \frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$
Mean Fractional Bias	$MFB = \frac{1}{N}\sum_{i=1}^{N}\frac{M_i - O_i}{(M_i + O_i)/2}$
Mean Fractional Error	$MFE = \frac{1}{N}\sum_{i=1}^{N}\frac{\lvert M_i - O_i\rvert}{(M_i + O_i)/2}$
RootMeanSquare Error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2}$
Ratio of Systematic and unsystematic RMSE	$RMSE_S \,/\, RMSE_U = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{M}_i - O_i)^2} \Big/ \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - \hat{M}_i)^2}$ *where* $\hat{M}_i = a + bO_i$ *are the regressed model values, estimated from a least square fit to observations;* $RMSE^2 = RMSE_S^2 + RMSE_U^2$.
Target	$RMSE \,/\, \sigma_O = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2} \Big/ \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - \overline{O})^2}$
Pearson Correlation Coefficient	$R = \sum_{i=1}^{N}(M_i - \overline{M})\cdot(O_i - \overline{O}) \Big/ \sqrt{\sum_{i=1}^{N}(M_i - \overline{M})^2} \cdot \sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}$
Index of Agreement	$IOA = 1 - N \cdot RMSE^2 \Big/ \sum_{i=1}^{N}\left(\lvert M_i - \overline{O}\rvert + \lvert O_i - \overline{O}\rvert\right)^2$
Relative Directive Error and its maximum	$RDE = \frac{\lvert O_{LV} - M_{LV}\rvert}{LV}$ *where* O_{LV} *is the closest observed concentration to the limit value concentration (LV) and* M_{LV} *is the correspondingly ranked modelled concentration.* *MRDE=Max (RDE over 90% of stations)*
Relative Percentile Error and its maximum	$RPE = \frac{\lvert O_p - M_p\rvert}{O_p}$ *where* p *is the percentile corresponding to the allowed number of exceedances of the limit value* *MRPE=Max (RPE over 90% of stations)*
Factor of modelled values within a factor of two of observations	$FAC2 = \frac{1}{N}\sum n_i \quad with \quad n_i = \begin{cases} 1 \text{ for } 0.5 \le \lvert M_i/O_i\rvert \le 2 \\ 0 \text{ else} \end{cases}$
Centred Root Mean Square error	$CRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[(M_i - \overline{M}) - (O_i - \overline{O})\right]^2}$
Model Efficiency Score	$MEF = 1 - RMSE^2$

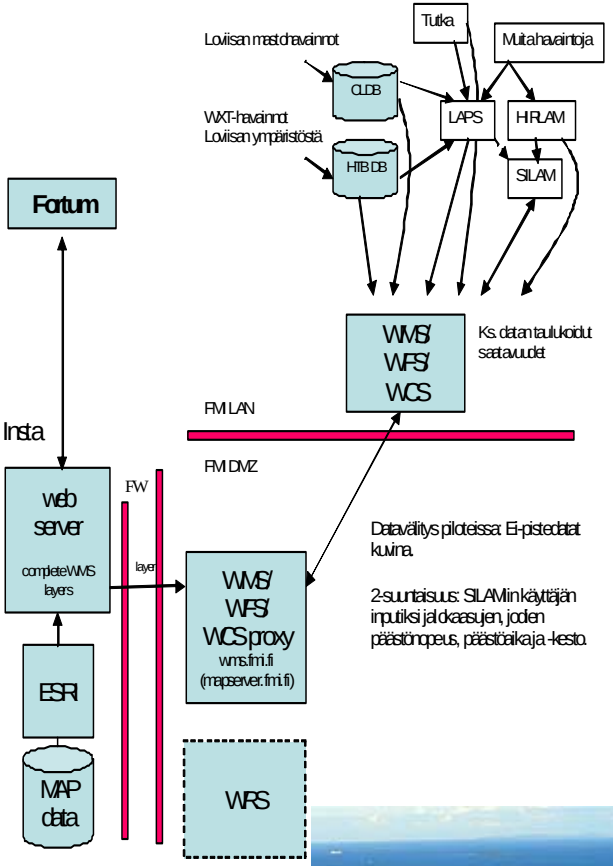
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

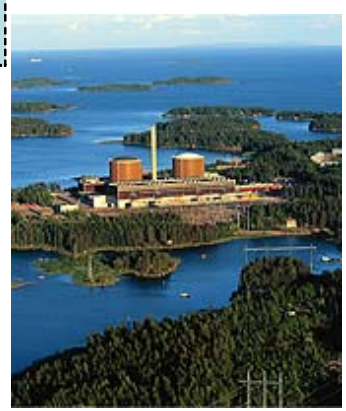
Air Quality / Practical applications



25.2.2008



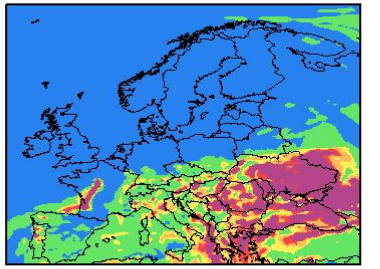




Ilmanlaatu nyt › Metsäpalot

På svenska | In English

Metsäpalosavujen leviämisennuste seuraaville 18 tunnille



09:00 28.08.
10:00 28.08.
11:00 28.08.
12:00 28.08.
13:00 28.08.
14:00 28.08.
15:00 28.08.
16:00 28.08.
17:00 28.08.
18:00 28.08.
19:00 28.08.
20:00 28.08.
21:00 28.08.
22:00 28.08.
23:00 28.08.
00:00 29.08.
01:00 29.08.
02:00 29.08.
PLAY

Hiukkaspitoisuus Pieni Kohtalainen Suuri

Kuvassa on mallilaskelmiin perustuva ennuste metsäpaloista leviävien pienhiukkasten pitoisuudelle lähituntien aikana. Vihreillä alueilla pitoisuus on pieni, keltaisilla ja oransseilla kohtalainen ja punaisilla suuri. Animaation voi keskeyttää siirtämällä osoittimen halutun kellonajan päälle. Animaatio lähtee uudelleen käyntiin kun osoitin siirretään PLAY-tekstin päälle.